

<https://helda.helsinki.fi>

Towards the completion of speciation : the evolution of reproductive isolation beyond the first barriers

Kulmuni, Jonna

2020-08-31

Kulmuni , J , Butlin , R K , Lucek , K , Savolainen , V & Westram , A M 2020 , ' Towards the completion of speciation : the evolution of reproductive isolation beyond the first barriers ' , Philosophical Transactions of the Royal Society. Biological Sciences , vol. 375 , no. 1806 , 20190528 . <https://doi.org/10.1098/rstb.2019.0528>

<http://hdl.handle.net/10138/336681>

<https://doi.org/10.1098/rstb.2019.0528>

cc_by

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Towards the completion of speciation: the evolution of reproductive isolation beyond the first barriers

Jonna Kulmuni¹, Roger K. Butlin^{2,3}, Kay Lucek⁴, Vincent Savolainen⁵ and Anja Marie Westram⁶

1. Organismal and Evolutionary Biology, University of Helsinki, Finland

2. Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

3. Department of Marine Sciences, University of Gothenburg, 405 30 Gothenburg, Sweden

4. Department of Environmental Sciences, University of Basel, 4056 Basel, Switzerland

5. Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK

6. IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

Abstract

Speciation, that is, the evolution of reproductive barriers eventually leading to complete isolation, is a crucial process generating biodiversity. Recent work has contributed much to our understanding of how reproductive barriers begin to evolve, and how they are maintained in the face of gene flow. However, little is known about the transition from partial to strong reproductive isolation (RI) and the completion of speciation. We argue that the evolution of strong RI is likely to involve different processes, or new interactions among processes, compared to the evolution of the first reproductive barriers. Transition to strong RI may be brought about by changing external conditions, for example following secondary contact. However, increasing levels of RI themselves create opportunities for new barriers to evolve, and for feedback, interaction or coupling among barriers. These changing processes may depend on genomic architecture and leave detectable signals in the genome. We outline outstanding questions and suggest more theoretical and empirical work, considering both patterns and processes associated with strong RI, is needed to understand how speciation is completed.

Is the evolution of strong reproductive isolation different from the evolution of the first barriers to gene flow?

Speciation is the evolution of reproductive isolation (RI) through the accumulation of barriers to gene exchange. Opinions vary on when the speciation process ends: some require RI to be complete (Barton and Hewitt 1985) while others allow some possibility for gene exchange (notably Coyne & Orr 2004, pp. 33-35); some require RI to be irreversible while others do not (Coyne and Orr, pp. 37-8). Divergence continues to accumulate after completion of RI. This is important for the evolution of biological diversity, as continued ecological divergence can facilitate co-existence in sympatry. Continued divergence can remove the possibility of speciation reversal when redundant barriers accumulate (Futuyma 1987). However, continuing divergence is not strictly part of the speciation process and in this Theme Issue we do not consider continuing divergence after the completion of RI.

Occasionally strong RI can appear rapidly, even in a single generation, in the case of ploidy change (Soltis, Soltis, and Tate 2003), chromosomal rearrangements (de Vos et al., this issue), and perhaps following hybridisation (Mallet 2007; Lamichhaney et al. 2018). However, more often the accumulation of barriers to gene flow is an extended process in which RI evolves slowly. Multiple different barriers evolve and are coupled (Butlin and Smadja 2018), typically over tens of thousands to millions of generations (Coyne and Orr 2004, Table 12.1). This suggests the idea of a 'speciation continuum', in which divergent populations can be placed according to their current level of RI. The continuum is not unidirectional nor unidimensional, but rather a trajectory in multivariate space (see also three-dimensional speciation cube in Dieckmann et al. 2004). During the speciation process, there may not be a monotonic increase in RI and the rate of accumulation of barriers to gene flow can vary widely (e.g. Nosil et al. 2017). Therefore, the speciation continuum is not a temporal progression: RI might stall at an intermediate level (Nosil, Harmon, and Seehausen 2009), it might decrease following a change in the environment that disrupts barriers to gene flow (e.g. Seehausen et al. 2008), or it might fluctuate over time.

In this Theme Issue we focus on the part of the speciation continuum where some barriers to gene flow have already evolved, as we are primarily concerned with the evolution of RI towards the completion of speciation. We contrast this with parts of the speciation process during which RI starts to accumulate and barriers to gene flow are being established against a background of weak or no divergence. Because RI rarely increases steadily with time, the contrast between parts of the continuum reflects the level of RI, not time since the start of a speciation event. Much attention in recent speciation research has focused on processes acting in parts of the continuum where RI is low, especially local adaptation. These studies are important but may not allow for inferences about the speciation process closer to completion. First, they do not increase our understanding of

the mechanisms and processes that can lead to the evolution of strong barriers. Second, the mix of processes contributing to speciation may change with increasing levels of RI and feedback between barriers and processes may occur. A key question then is, in what way are the contributing barriers, processes and genomic patterns different between the different parts of the speciation continuum?

Ideally, to address questions on the evolution of strong RI one would sample population pairs across the continuum and test whether the increase in RI is consistent with a specific process (e.g. a “snowballing” of hybrid incompatibilities (Orr 1995)); or whether patterns associated with a certain process increase along the continuum (e.g. signatures of reinforcement). Even when it is not possible to analyse comparable population pairs across the speciation continuum within taxa, we can learn from individual case studies of taxon pairs near the completion of speciation. These studies will contribute to an understanding of the reproductive barriers and genomic patterns that characterise the evolution of strong RI (Stankowski et al. and Yamasaki et al. this issue) and, when combined, can give important hints about the underlying processes. Theoretical studies and simulation approaches are necessary to complement empirical work and will allow for a more direct understanding of processes facilitating strong RI (Blanckaert et al., Payne and Polechova and Bisschop et al. this issue). This Theme Issue combines theoretical and empirical work on the progression to strong RI and considers the patterns and processes associated with high levels of isolation from various angles. We hope it will contribute to a more comprehensive view and a deeper understanding of the speciation process.

Are there stages in the speciation process and what causes them?

General evolutionary processes that are known to contribute to speciation include differential adaptation driven by divergent selection, divergence in mating traits due to sexual selection, and mutation-order processes resulting in incompatibilities driven by either selection or drift, the mix of effects being dependent on spatial context (Coyne and Orr 2004). These processes can lead to the evolution of extrinsic barriers (divergent adaptation) as well as intrinsic barriers and assortment. Importantly, these processes can take place in any part of a speciation trajectory. When some RI is already present, they may strengthen existing barriers that evolved earlier, lead to the emergence of new barriers, or increase RI by coupling of barriers (i.e. by causing the effects of multiple barriers to work together in opposing gene flow; Butlin and Smadja 2018).

There is a long-standing idea that the continuum might be divided into more or less discrete ‘stages’ or ‘phases’ (here we use the former). This idea dates back to Darwin and Wallace but has received renewed attention recently (see Lowry (2012) for a history of the debate over ‘stages of speciation’). Hendry et al. (2009) and Mérot et al. (2017) illustrate the heuristic value of this view

when applied to empirical cases. If separate stages exist, they might reflect changing patterns and processes along the speciation continuum; therefore, identifying and understanding what causes stages will directly relate to the questions posed by this Theme Issue.

Changes in the strength of RI, the rate at which RI accumulates or the forces driving the speciation process could all be used to define stages. The stages might also be imposed externally. The classic example is divergence in allopatry followed by secondary contact and hybridization, which is expected to introduce new potential mechanisms for the evolution of reproductive isolation, notably reinforcement. Another example of stages associated with external factors is described by Muschick et al. (this issue). They show that shifts between closely-related host plants generate little RI (in terms of host preference), while shifts between more distantly-related hosts occur rarely, but lead to greater RI in *Timema* stick insects. Host switches could therefore help populations move to a stage with stronger RI.

Dieckmann et al. (2004) argued for the necessity of recognising stages in generating a coherent overview of speciation precisely because of the changes to the mix of individual processes involved. Others have argued that the initiation of RI, its accumulation and its completion are likely to involve different mixes of processes, different impacts of the spatial setting and of the genetic architecture (Butlin, Galindo, and Grahame 2008). Focusing on genetic architecture, and on speciation with gene flow, Wu (2001), Feder, Egan, and Nosil (2012) and Feder et al. (2012) considered the existence of stages. They suggested the initial stage is dominated by direct selection on individual barrier loci leading to localised genomic barriers, which then extend around directly-selected loci due to physical linkage, and finally to a more genome-wide barrier effect leading towards complete RI. Stages with higher RI have a stronger contribution of indirect selection via linkage disequilibrium (LD). Similar transitions between regimes dominated by direct selection vs. LD have been described in earlier work (Barton 1983). Feder et al. (2012) suggested that the build-up of LD might generate a sharp transition between these regimes under divergence with gene flow. If so, there may be very distinct stages in some speciation trajectories.

Under which conditions we expect distinct stages remains an open question (Box 1: Question 1). In addition, in order to better understand speciation, it is not sufficient to identify stages (if they exist). Rather, we need to ask whether, how and why the processes contributing to speciation change along the speciation continuum.

The dynamics of RI accumulation are likely to change with increasing RI

In principle, barriers resulting from the basic processes described above could simply accumulate and steadily increase reproductive isolation. Species would just be divergent populations “writ

large", as suggested for *Rhagoletis* flies (Powell et al. 2013; Meyers et al., this issue). Nosil et al. (2017) argue that, even with gene flow, strong selection on individual loci might generate this kind of steady accumulation of RI (in contrast to weaker selection on multiple loci which can create sharp transitions, see below).

Such a steady accumulation of RI seems unlikely in most scenarios, due to the emergence of new processes along the speciation continuum (discussed below). These may lead to feedback or interaction between different barriers and processes, making these emergent properties important factors in the evolution of strong RI (Box 1: Question 2). Even without changing processes and mechanisms there are reasons to expect differences between parts of the speciation continuum. For example, if some types of barriers evolve systematically more slowly (e.g. because they require large numbers of new mutations), they are likely to appear later in the speciation process. Even if this is not the case and different barriers emerge completely at random, one could expect the relative contribution of postzygotic barriers to RI to decrease over the course of speciation. This is simply because it becomes more likely that the emergence of barriers acting earlier in the life cycle decreases the impact of late-acting barriers on total isolation. This argument is general (Turelli, Barton, and Coyne 2001); it does not rely on reinforcement of prezygotic barriers.

Although possible, and perhaps useful as a null model, speciation trajectories that are not characterised by feedback mechanisms or interplay between different barriers as RI increases are probably rare in nature. This is because the introduction of barriers to gene flow has many effects: LD is increased, particularly between loci underlying RI, effective gene flow is reduced, there is stronger selection against hybrids, and a larger potential for negative epistatic interactions between alleles from different populations (Orr 1995). In addition, once some initial barriers have established coupling between pre- and postzygotic barriers (Barton and de Cara 2009), intrinsic and extrinsic barriers (Bierne et al. 2011) or between different barrier loci (Barton 1983) can occur. All these features can modify the trajectory of speciation, potentially leading to different barrier types, effect sizes, and genomic patterns at higher levels of RI compared to low RI. Add to this the likelihood of externally-imposed changes in geographical distribution and demography, and uniform speciation trajectories become very improbable (e.g. cessation of gene flow by strong RI in cichlids requires a combination of premating prezygotic isolation supplemented with intrinsic and extrinsic postzygotic barriers whose strengths and importance vary among different adaptive radiations; Rometsch et al. this issue). Some changes will not occur in every speciation trajectory (e.g. addition of reinforcement) while others might be inevitable (e.g. accumulation of incompatibilities in allopatry; (Orr 1995)).

Below, we discuss feedback processes and mechanisms that lead to coupling of barrier effects, especially reinforcement, that can only occur once some RI already exists. These are likely to be

the dominant reasons for changes in the processes operating in different parts of the speciation continuum.

Feedback, interactions and the snowball effect

Processes that drive divergence (e.g. divergent selection, sexual selection, drift) and increase RI may vary along the speciation continuum, e.g. because LD increases, effective gene flow decreases (unless in complete allopatry), and the total number of barrier loci increases with higher RI. Both in allopatry and in settings with gene flow these changes can lead to positive feedback, altering the rate at which RI increases. Positive feedback, where some initial barrier allows the spreading of the barrier effect or increases opportunities for other barriers to establish, might generate rapid transitions from weak to strong overall isolation and so justify the identification of distinct stages in the speciation process. Feedback within and between processes can be hard to isolate in natural populations but separation can be achieved in models.

One type of feedback acting in the presence of gene flow is divergence hitchhiking, where a divergently selected locus provides a barrier locally in the genome that allows accumulation of differentiation at linked sites, while gene flow continues to homogenise allele frequencies elsewhere in the genome (Via and West 2008). Feedback occurs if the localised barrier facilitates the spread of other barrier alleles allowing the size of the 'genomic island of differentiation' and the total RI to increase. The efficacy of this process depends on the extent of LD and so the chance that new mutations will benefit from the reduction in gene flow. Its effectiveness has been debated (Feder and Nosil 2010; Rafajlović et al. 2016; Aeschbacher and Bürger 2014). Arguably, divergence hitchhiking is mostly relevant when total RI is still low, because with strong RI the density of selected loci becomes high enough to cause LD even between distant and physically unlinked loci, generating a general barrier to gene flow. 'Genome hitchhiking' (a genome-wide reduction in the effective migration rate) could then be considered characteristic of situations closer to complete RI (Feder et al. 2014) (Box 1: Question 4).

The interaction between selection and recombination can generate a sharp discontinuity in the strength of the genome-wide barrier to gene flow (e.g. *Populus* trees in Shang et al. and sticklebacks in Yamasaki et al., this issue). In the context of secondary contact, Barton (1983) showed that a strong genome-wide barrier is expected when the ratio of selection to recombination (known as the 'coupling coefficient') exceeds a critical value due to accumulation of incompatibilities (see also Barton and De Cara 2009). Above this critical value, secondary contact is much more likely to lead to speciation; for example, because the associations needed for reinforcement to be effective are much more likely to be maintained. An analogous transition is seen in the models of speciation with continuous gene flow by Flaxman et al. (2014): where the

genomic architecture allows for LD (i.e. in models where sets of genes are transmitted from parent to offspring, with or without linkage, rather than sampled randomly from the population) and where selection per locus is weak relative to migration ($s \ll m$), reproductive isolation accumulates slowly until the genomic density of selected loci reaches a critical point where LD can build up. LD increases the effective selection on each locus and this generates positive feedback leading to a rapid increase in RI. However, this result is based on the assumption of locally adaptive mutations under weak selection, and modifications to the model might change whether and how rapidly the feedback process occurs. For example, the occurrence of globally beneficial mutations is excluded from Flaxman et al. (2014) and other similar models of local adaptation. When they occur, their sweeps interfere with the evolution of local adaptation, reducing the differentiation between populations and delaying the evolution of strong RI (Bisschop et al., this issue).

Nosil et al. (2017; p.1) suggested that behaviours analogous to those seen in the Flaxman et al. (2014) models may be more general: “divergence process[es] involving small changes can suddenly speed up at a ‘tipping point’ in speciation, at least in theory”. How widespread these effects are, in theory or in reality, is currently an open question (Box 1: Question 3 and 4). There is some empirical evidence consistent with tipping points (Nosil et al. 2017). The ‘grey zone’ of speciation studied by Roux et al. (2016), shows a transition zone between 0.5% and 2% net synonymous divergence during which the probability of gene flow decreases sharply, which could be interpreted as consistent with the tipping point idea. On the other hand, the identified ‘grey zone’ is quite wide and contains multiple semi-isolated taxa.

In addition to hitchhiking, the accumulation of genetic incompatibilities can accelerate the evolution of RI, and so generate marked changes across the speciation continuum. Coughlan & Matute (this issue) argue that genetic incompatibilities may be important already in the beginning of the speciation process and discuss how they can enhance the evolution of strong reproductive isolation via reinforcement and the ‘snowball effect’. This effect occurs because the number of genetic incompatibilities increases faster than linearly with time if each new mutation has the potential to be incompatible with any other mutation (Orr 1995). In Orr’s model, the number of potential pairwise incompatibilities increases with the square of the number of genetic differences that have accumulated between two populations. This acceleration potentially operates in all spatial settings. Evidence for the snowball effect is mixed (Gourbière and Mallet 2010; Presgraves 2010; Matute et al. 2010; Moyle and Nakazato 2010; Guerrero et al. 2017). Importantly, the rate at which incompatibilities accumulate, and so the snowball effect, depends on the number and distribution of epistatic interactions in the genome. Models explicitly incorporating empirical gene interaction networks might lead to different predictions than standard population genetic models, as discussed by Satokangas et al. (this issue). As more data have accumulated on the molecular basis of epistatic interactions within the genome, models of incompatibilities have incorporated

more realistic assumptions. Kalirad and Azevedo (2017) and Dagilis, Kirkpatrick, and Bolnick (2019) showed that considering both positive and negative epistasis and accounting for differences in within-population and between-population epistasis predicts a snowball-like effect for hybrid fitness. In these models, loss of hybrid fitness can accelerate due to changes in the magnitude and/or direction of average epistatic effects of mutations over time, rather than changes in the numbers of incompatibilities. On the other hand, Blanckaert and Hermisson (2018) have shown that pre-existing incompatibilities do not always facilitate invasion of further barrier genes under gene flow, which can seriously disrupt the snowball effect. Clearly, open questions remain (Box 1: Question 5).

A different feedback process occurs when each evolutionary change can open up unique new possibilities. As two populations diverge at functional loci, the set of new mutations that are positively selected in one population will start to diverge from the set that is positively selected in the other population. This is partly due to direct epistatic interactions between mutations and accumulating structural variation in the genome (Navarro & Barton 2003) and partly due to a changing ecological environment enabled by divergence. It is related to mutation-order effects (Schluter 2009) in the sense that the events governing the first mutations to spread in each population set them off on different trajectories. This is another possible area for future research (Box 1: Question 6). One could argue that reinforcement falls into this category: the spread of alleles that reduce hybrid fitness creates a selection pressure for assortative mating that was initially absent (Garner et al. 2018). However, we treat reinforcement separately (see following section).

Remarkably, a strengthening of RI might occur even if existing barriers enable the spread of *the same* new mutation in both diverging populations. In their model, Blanckaert et al. (this issue) show that an incompatibility between two derived alleles, with at least one of them being locally adapted, can generate a much stronger barrier if it is accompanied by the fixation of the same new allele at a third locus, if this allele displays a specific set of epistatic interactions.

Coupling and reinforcement

The barriers that form in the initial stages of speciation may not coincide, in the sense that they might separate different groups of populations in space or in some other dimension (such as time or niche). For example, local adaptation might distinguish groups of populations occupying different habitats while incompatible allelic combinations separate groups of populations with a common demographic history. Such spatially scattered barriers may often be hard to detect empirically, as they are often weak and affect only a small number of loci. RI can be enhanced by processes that bring these barrier effects together so that they separate coincident groups of

populations ('coupling' *sensu* Butlin and Smadja 2018). Spatial movement of clines that attract one another when they overlap (Bierne et al. 2011), or cycles of population expansion and contraction (Hewitt 1989) are examples of potential processes that can achieve coupling of barrier effects after the evolution of some RI, i.e. in the part of the speciation continuum where substantial barriers already exist. A similar effect might be caused by short periods of allopatry because divergence at loci contributing to local adaptation is limited by gene flow, particularly if locally adapted alleles also cause some incompatibility in hybrids.

In addition to coupling of existing barrier effects, the presence of some initial barriers may result in selection favouring the evolution of new barrier effects or the enhancement of existing effects. This form of coupling includes reinforcement as traditionally viewed (indirect selection favouring an increase in prezygotic isolation as a result of reduced fitness of hybrids), but also other processes where indirect selection contributes to the origin or strengthening of barriers, whether pre- or postzygotic (Butlin and Smadja 2018). All of these processes are most likely to operate when the initial reproductive isolation is already strong. They might be supplemented by direct selection pressures, such as the costs of wasted mating efforts or sexual selection induced by changes in female preference (Sachdeva and Barton 2017). Tinghitella et al. (this issue) further introduce an ecological element, implicating different ecological conditions that change mating interactions and select for environment-dependent mate preferences. Payne and Polechova (this issue) argue that pre-existing niche preference speeds up further increase of assortment, thus facilitating reinforcement in the later stages of speciation.

The classical scenario for reinforcement is the enhancement of assortative mating following secondary contact. The spatial distributions of most taxa have fluctuated through time, with many populations temporarily persisting in geographical isolation from each other, for example in distinct glacial refugia (Hewitt 2000). Speciation can then have two clear stages: during the period of allopatry populations accumulate intrinsic and/or extrinsic incompatibilities and these postzygotic barriers act upon secondary contact to drive an enhancement of isolation by reinforcement. Both modelling and empirical data now strongly support the possibility of reinforcement in this scenario, although the outcome depends on many factors (Servedio and Noor 2003). Strong post-zygotic isolation, multiple-effect traits (Smadja and Butlin 2011), low costs of mate choice, matching mechanisms of assortative mating (Kopp et al. 2018) and one-allele systems (where no divergence between populations is required to enhance assortment; Felsenstein 1981; Servedio 2000) all promote the evolution of assortative mating. Although the possibility of classical reinforcement is now well established, its overall contribution to speciation remains unclear (Box 1: Question 7). There have been few attempts to measure its frequency (Yukilevich 2012). Although patterns of reinforcement are often observed at the phenotypic level, the genomic footprint of reinforcement is much less well understood and data are scarce (Ortiz-Barrientos, Counterman, and Noor 2004;

Garner et al. 2018) (Box 1: Question 7). Focusing on copy-number variation detectable in paired-end sequencing data, North et al. (this issue) test for the expected patterns of genomic divergence under a reinforcement scenario and show that these common, but understudied fine-scale structural variants make for strong candidate targets of reinforcing selection. While some ‘adaptive coupling’ processes have been well-studied (especially classical reinforcement), others have not (Box 1: Question 7). For example, habitat choice is a potent source of assortment but it is a relatively under-studied form of barrier (Edelaar and Bolnick 2012; Berner and Thibert-Plante 2015). However, these processes might also be self-limiting in the sense that the selection pressure for increased isolation derives from the production of hybrids and so declines as hybrids become rare, for whatever reason (Yukilevich and True 2006). Self-limitation might also apply to other processes that mainly occur when RI is already relatively strong, for example fitness gains from approaching adaptive optima decline as the optima are approached, potentially leading to a stable balance between divergent selection and gene flow (Nosil, Harmon, and Seehausen 2009; Servedio and Hermisson 2019).

The evolution or enhancement of new barriers to gene flow as a result of costs associated with existing barriers (e.g. Turner, Jacobson, and Taylor 2010) is a much more general feature of systems with high RI than suggested by the historical focus on the classical form of reinforcement (Butlin and Smadja 2018). These processes can only operate when populations are in contact because costs of hybridisation are not expressed during periods of allopatry, and only to a limited degree in narrow hybrid zones. These mechanisms then require some minimum amount of interbreed. However, they do not require divergence followed by secondary contact: they are also potentially a key part of speciation with continuous gene flow (e.g. Dieckmann and Doebeli 1999, Osborne et al., this issue). They are not limited to assortative mating, or even to enhancement of pre-zygotic isolation (Butlin and Smadja 2018). More work is needed on these other forms of coupling (Box 1: Question 7).

Conclusion

We expect the parts of the speciation continuum where substantial RI is already in place to have distinctive features and contain a different mix of processes, including feedback between barriers, compared to initiation of RI. We suggest that more attention to these features and processes is important for a full understanding of speciation, and we present a list of open questions in Box 1. Natural systems, but also analysis of anthropogenic and experimental hybridization events, can provide clues to answering open questions (e.g. Viard, Riginos and Bierne, this issue).

Empirically understanding how processes change along the speciation continuum is challenging, particularly because the necessary contrasts are not available in most taxa, and because many of

the processes are notoriously difficult to study even for a single taxon pair (e.g. reinforcement). Progress will come from a combination of case studies of taxon pairs nearing complete isolation (e.g. Stankowski et al., Yamasaki et al., Osborne et al. and North et al. this issue), comparative approaches (e.g. de Vos et al. and Meyers et al. this issue), theoretical studies highlighting what drives and hinders the evolution of strong RI (e.g. Blanckaert et al., Bisschop et al. and Payne & Polechova, this issue), meta-analyses (Rometsch et al. and Shang et al. this issue), the identification of processes defining stages (Tinghitella et al. and Muschick et al. this issue) and combined approaches as outlined in Coughlan & Matute (this issue) and Satokangas et al. (this issue).

Box 1. Outstanding questions about the progression towards complete RI during speciation.

1. What underlying processes create heterogeneity and the appearance of stages in the speciation continuum (e.g. Muschick et al., this issue)? Do distinct stages occur in some speciation trajectories and not others?
2. Are there systematic differences in the types of barriers present at different points in the speciation continuum (Coyne & Orr 2004)? The classical expectation is for postzygotic barriers (local adaptation or intrinsic incompatibility) to evolve first, followed by prezygotic barriers (in a trajectory where allopatric divergence is followed by secondary contact, e.g. North et al., this issue) but alternative scenarios have been proposed (Coughlan and Matute, this issue). For example, learned behavioural barriers could provide the reduction in gene flow needed for the establishment of ecological differentiation (e.g. Rometsch et al., this issue). Are changes in barrier types along the speciation continuum driven mainly by external conditions (such as secondary contact) or by opportunities created by the first barriers to appear (e.g. Tinghitella et al., this issue)? Do barriers evolve independently or in concert?
3. Are some speciation trajectories characterised by tipping points, or bi-stable states, resulting in sharp transitions from slow to rapid accumulation of RI? If so, does rapid accumulation lead directly to completion of speciation or is it followed by a slower stage involving other processes (such as reinforcement)? What is the evidence for tipping points and what mechanisms contribute to progressing through tipping points?
4. Under what circumstances do divergence and genomic hitchhiking contribute to the accumulation of RI, when in the speciation continuum do they contribute, and can they be

used to distinguish between different stages of speciation? What is the empirical evidence for these processes and how can the existing evidence be extended and improved? Does moving from 2-deme models to more realistic models of continuous space lead to different predictions for the role of divergence and genomic hitchhiking?

5. The snowball effect depends on assumptions of epistatic interactions in the genome (Satokangas et al., this issue). Does accumulation of incompatibilities accelerate in snowball fashion under biologically relevant assumptions of epistasis? Are complex incompatibilities involving multiple loci more stable in the face of gene flow (Blanckaert et al., this issue) and so important for the completion of speciation compared to 2-locus incompatibilities? How common are incompatibilities whose expression depends on the environment compared to purely intrinsic incompatibilities?
6. Any change in the genetic composition of a population alters the opportunities and constraints on future evolution. How much does this kind of feedback contribute to divergence between populations and the accumulation of RI? Does this occur mainly by stochastic 'mutation order' processes, co-evolution of gene networks, new ecological opportunities opened up by early adaptive substitutions, or by processes such as sexual or genomic conflict or sexual selection that are partly independent of the environment?
7. Reinforcement is a classical process in cases where some RI already exists, but it remains unclear how often, and how much it contributes to speciation (North et al., this issue). Does reinforcement complete RI following secondary contact? If so, is this only when the initial isolation is already strong? How common is reinforcement in speciation trajectories with continuous or frequent gene flow? How widely does reinforcement in the broad sense (adaptive coupling with barrier enhancement, Butlin and Smadja 2018) operate across different types of barrier effects?

Acknowledgements

This Theme Issue is derived from a symposium on "Late stages in speciation: evolution of strong reproductive isolation in the presence of gene flow" organised by the guest editors at the II Joint Congress on Evolutionary Biology in Montpellier, France, in August 2018. We are grateful for all the authors of the Theme Issue who provided comments on our manuscript. JK was supported by Academy of Finland grant no. 328961 and European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 750015. RKB was supported by ERC and NERC. AMW was funded by the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 797747. VS was supported by NERC.

Author contributions

JK, AMW, VS, KL and RKB conceived the framework of the paper. RKB wrote the first draft with contributions from JK, AMW and KL. JK revised the manuscript based on external comments. All authors critically reviewed the manuscript and accepted the final version.

References

- Aeschbacher, Simon, and Reinhard Bürger. 2014. "The Effect of Linkage on Establishment and Survival of Locally Beneficial Mutations." *Genetics* 197 (1): 317–36. <https://doi.org/10.1534/genetics.114.163477>.
- Barton, N. H. 1983. "Multilocus Clines." *Evolution* 37 (3): 454–71. <https://doi.org/10.2307/2408260>.
- Barton, N H, and G M Hewitt. 1985. "Analysis of Hybrid Zones." *Annual Review of Ecology and Systematics* 16 (1): 113–48. <https://doi.org/10.1146/annurev.es.16.110185.000553>.
- Barton, Nicholas H., and Maria Angeles Rodriguez De Cara. 2009. "The Evolution of Strong Reproductive Isolation." *Evolution* 63 (5): 1171–1190. <https://doi.org/10.1111/j.1558-5646.2009.00622.x>.
- Berner, D., and X. Thibert-Plante. 2015. "How Mechanisms of Habitat Preference Evolve and Promote Divergence with Gene Flow." *Journal of Evolutionary Biology* 28 (9): 1641–55. <https://doi.org/10.1111/jeb.12683>.
- Bierne, Nicolas, John Welch, Etienne Loire, François Bonhomme, and Patrice David. 2011. "The Coupling Hypothesis: Why Genome Scans May Fail to Map Local Adaptation Genes." *Molecular Ecology* 20 (10): 2044–72. <https://doi.org/10.1111/j.1365-294X.2011.05080.x>.
- Blanckaert, Alexandre, and Joachim Hermisson. 2018. "The Limits to Parapatric Speciation II: Strengthening a Preexisting Genetic Barrier to Gene Flow in Parapatry." *Genetics* 209 (1): 241–54. <https://doi.org/10.1534/genetics.117.300652>.
- Butlin, Roger K, Juan Galindo, and John W Grahame. 2008. "Sympatric, Parapatric or Allopatric: The Most Important Way to Classify Speciation?" *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1506): 2997–3007. <https://doi.org/10.1098/rstb.2008.0076>.
- Butlin, Roger K, and Carole M Smadja. 2018. "Coupling , Reinforcement , and Speciation" 191 (2): 155–72. <https://doi.org/10.1086/695136>.
- Coyne, Jerry A., and H. Allen Orr. 2004. *Speciation*. Sinauer Associates.
- Dagilis, Andrius J., Mark Kirkpatrick, and Daniel I. Bolnick. 2019. "The Evolution of Hybrid Fitness during Speciation." *PLOS Genetics* 15 (5): e1008125. <https://doi.org/10.1371/journal.pgen.1008125>.
- Dieckmann, Ulf, and Michael Doebeli. 1999. "On the Origin of Species by Sympatric Speciation." *Nature* 400 (6742): 354–57. <https://doi.org/10.1038/22521>.

- Dieckmann, Ulf, Michael Doebeli, Johan A.J. Metz, and Diethard Tautz. 2012. "Adaptive Speciation." *Adaptive Speciation*, 1–460. <https://doi.org/10.1017/CBO9781139342179>.
- Edelaar, Pim, and Daniel I. Bolnick. 2012. "Non-Random Gene Flow: An Underappreciated Force in Evolution and Ecology." *Trends in Ecology & Evolution* 27 (12): 659–65. <https://doi.org/10.1016/j.tree.2012.07.009>.
- Feder, Jeffrey L., and Patrik Nosil. 2010. "THE EFFICACY OF DIVERGENCE HITCHHIKING IN GENERATING GENOMIC ISLANDS DURING ECOLOGICAL SPECIATION." *Evolution* 64 (6): 1729–47. <https://doi.org/10.1111/j.1558-5646.2010.00943.x>.
- Feder, Jeffrey L., Patrik Nosil, Aaron C. Wacholder, Scott P. Egan, Stewart H. Berlocher, and Samuel M. Flaxman. 2014. "Genome-Wide Congealing and Rapid Transitions across the Speciation Continuum during Speciation with Gene Flow." *Journal of Heredity* 105 (S1): 810–20. <https://doi.org/10.1093/jhered/esu038>.
- Feder, Jeffrey L., Scott P. Egan, and Patrik Nosil. 2012. "The Genomics of Speciation-with-Gene-Flow." *Trends in Genetics : TIG* 28 (7): 342–50. <https://doi.org/10.1016/j.tig.2012.03.009>.
- Feder, Jeffrey L., Richard Gejji, Sam Yeaman, and Patrik Nosil. 2012. "Establishment of New Mutations under Divergence and Genome Hitchhiking." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 367 (1587): 461–74. <https://doi.org/10.1098/rstb.2011.0256>.
- Felsenstein, Joseph. 1981. "Skepticism towards Santa Rosalia, or Why Are There so Few Kinds of Animals?" *Evolution* 35 (1): 124–38. <https://doi.org/10.2307/2407946>.
- Flaxman, Samuel M., Aaron C. Wacholder, Jeffrey L. Feder, and Patrik Nosil. 2014. "Theoretical Models of the Influence of Genomic Architecture on the Dynamics of Speciation." *Molecular Ecology* 23 (16): 4074–88. <https://doi.org/10.1111/mec.12750>.
- Garner, Austin, Benjamin Goulet, Matthew Farnitano, Y. Molina-Henao, and Robin Hopkins. 2018. "Genomic Signatures of Reinforcement." *Genes* 9 (4): 191. <https://doi.org/10.3390/genes9040191>.
- Gourbière, Sébastien, and James Mallet. 2010. "Are Species Real? The Shape of the Species Boundary with Exponential Failure, Reinforcement, and the 'Missing Snowball.'" *Evolution* 64 (1): 1–24. <https://doi.org/10.1111/j.1558-5646.2009.00844.x>.
- Guerrero, Rafael F., Christopher D. Muir, Sarah Josway, and Leonie C. Moyle. 2017. "Pervasive Antagonistic Interactions among Hybrid Incompatibility Loci," 1–19.
- Hendry, A. P., D. I. Bolnick, D. Berner, and C. L. Peichel. 2009. "Along the Speciation Continuum

- in Sticklebacks.” *Journal of Fish Biology* 75 (8): 2000–2036. <https://doi.org/10.1111/j.1095-8649.2009.02419.x>.
- Hewitt, Godfrey. 2000. “The Genetic Legacy of the Quaternary Ice Ages.” *Nature* 405 (6789): 907–13. <https://doi.org/10.1038/35016000>.
- Kalirad, Ata, and Ricardo B R Azevedo. 2017. “Spiraling Complexity: A Test of the Snowball Effect in a Computational Model of RNA Folding.” *Genetics* 206 (1): 377–88. <https://doi.org/10.1534/genetics.116.196030>.
- Kopp, Michael, Maria R. Servedio, Tamra C. Mendelson, Rebecca J. Safran, Rafael L. Rodríguez, Mark E. Hauber, Elizabeth C. Scordato, et al. 2018. “Mechanisms of Assortative Mating in Speciation with Gene Flow: Connecting Theory and Empirical Research.” *The American Naturalist* 191 (1): 1–20. <https://doi.org/10.1086/694889>.
- Lamichhaney, Sangeet, Fan Han, Matthew T Webster, Leif Andersson, B Rosemary Grant, and Peter R Grant. 2018. “Rapid Hybrid Speciation in Darwin’s Finches.” *Science (New York, N.Y.)* 359 (6372): 224–28. <https://doi.org/10.1126/science.aao4593>.
- Lowry, David B. 2012. “Ecotypes and the Controversy over Stages in the Formation of New Species.” *Biological Journal of the Linnean Society* 106 (2): 241–57. <https://doi.org/10.1111/j.1095-8312.2012.01867.x>.
- Mallet, James. 2007. “Hybrid Speciation.” *Nature* 446 (7133): 279–83. <https://doi.org/10.1038/nature05706>.
- Matute, D. R., I. A. Butler, D. A. Turissini, and J. A. Coyne. 2010. “A Test of the Snowball Theory for the Rate of Evolution of Hybrid Incompatibilities.” *Science* 329 (5998): 1518–21. <https://doi.org/10.1126/science.1193440>.
- Mérot, C., C. Salazar, R. M. Merrill, C. D. Jiggins, and M. Joron. 2017. “What Shapes the Continuum of Reproductive Isolation? Lessons from *Heliconius* Butterflies.” *Proceedings of the Royal Society B: Biological Sciences* 284 (1856). <https://doi.org/10.1098/rspb.2017.0335>.
- Moyle, Leonie C, and Takuya Nakazato. 2010. “Hybrid Incompatibility “Snowballs” between *Solanum* Species.” *Science (New York, N.Y.)* 329 (5998): 1521–23. <https://doi.org/10.1126/science.1193063>.
- Nosil, Patrik, Jeffrey L. Feder, Samuel M. Flaxman, and Zachariah Gompert. 2017. “Tipping Points in the Dynamics of Speciation.” *Nature Ecology & Evolution* 1 (2): 0001. <https://doi.org/10.1038/s41559-016-0001>.
- Nosil, Patrik, Luke J. Harmon, and Ole Seehausen. 2009. “Ecological Explanations for

- (Incomplete) Speciation.” *Trends in Ecology & Evolution* 24 (3): 145–56.
<https://doi.org/10.1016/j.tree.2008.10.011>.
- Orr, H. A. 1995. “The Population Genetics of Speciation: The Evolution of Hybrid Incompatibilities.” *Genetics* 139 (4): 1805–13.
- Ortiz-Barrientos, Daniel, Brian A Counterman, and Mohamed A. F Noor. 2004. “The Genetics of Speciation by Reinforcement.” Edited by Nick Barton. *PLoS Biology* 2 (12): e416.
<https://doi.org/10.1371/journal.pbio.0020416>.
- Powell, Thomas H.Q., Glen R. Hood, Mason O. Murphy, Jeffrey S. Heilveil, Stewart H. Berlocher, Patrik Nosil, and Jeffrey L. Feder. 2013. “Genetic Divergence along the Speciation Continuum: The Transition from Host Race to Species in *Rhagoletis* (Diptera: Tephritidae).” *Evolution* 67 (9): 2561–76. <https://doi.org/10.1111/evo.12209>.
- Presgraves, Daven C. 2010. “Speciation Genetics: Search for the Missing Snowball.” *Current Biology* 20 (24): R1073–74. <https://doi.org/10.1016/J.CUB.2010.10.056>.
- Rafajlović, Marina, Anna Emanuelsson, Kerstin Johannesson, Roger K. Butlin, and Bernhard Mehlig. 2016. “A Universal Mechanism Generating Clusters of Differentiated Loci during Divergence-with-Migration.” *Evolution* 70 (7): 1609–21. <https://doi.org/10.1111/evo.12957>.
- Sachdeva, Himani, and Nicholas H. Barton. 2017. “Divergence and Evolution of Assortative Mating in a Polygenic Trait Model of Speciation with Gene Flow.” *Evolution* 71 (6): 1478–93.
<https://doi.org/10.1111/evo.13252>.
- Schluter, D. 2009. “Evidence for Ecological Speciation and Its Alternative.” *Science* 323 (5915): 737–41. <https://doi.org/10.1126/science.1160006>.
- Seehausen, Ole, Gaku Takimoto, Denis Roy, and Jukka Jokela. 2008. “Speciation Reversal and Biodiversity Dynamics with Hybridization in Changing Environments.” *Molecular Ecology* 17 (1): 30–44. <https://doi.org/10.1111/j.1365-294X.2007.03529.x>.
- Servedio, Maria R. 2000. “REINFORCEMENT AND THE GENETICS OF NONRANDOM MATING.” *Evolution* 54 (1): 21–29. <https://doi.org/10.1111/j.0014-3820.2000.tb00003.x>.
- Servedio, Maria R., and Joachim Hermisson. 2019. “The Evolution of Partial Reproductive Isolation as an Adaptive Optimum.” *Evolution* 74 (1): 4–14. <https://doi.org/10.1111/evo.13880>.
- Servedio, Maria R., and Mohamed A.F. Noor. 2003. “The Role of Reinforcement in Speciation: Theory and Data.” *Annual Review of Ecology, Evolution, and Systematics* 34 (1): 339–64.
<https://doi.org/10.1146/annurev.ecolsys.34.011802.132412>.
- Smadja, Carole M., and Roger K. Butlin. 2011. “A Framework for Comparing Processes of

Speciation in the Presence of Gene Flow.” *Molecular Ecology* 20 (24): 5123–40.
<https://doi.org/10.1111/j.1365-294X.2011.05350.x>.

Society, The American, and Chicago Press. 2016. “On the Role of Species in Anagenesis Author (s): Douglas J. Futuyma Source: The American Naturalist, Vol. 130, No. 3 (Sep., 1987), Pp. 465–473 Published by: The University of Chicago Press for The American Society of Naturalists Stable URL: ” 130 (3): 465–73.

Soltis, Douglas E., Pamela S. Soltis, and Jennifer A. Tate. 2003. “Advances in the Study of Polyploidy since Plant Speciation.” *New Phytologist* 161 (1): 173–91.
<https://doi.org/10.1046/j.1469-8137.2003.00948.x>.

Turelli, M., N. H. Barton, and J. A. Coyne. 2001. “Theory and Speciation.” *Trends in Ecology & Evolution* 16 (7): 330–43. [https://doi.org/10.1016/s0169-5347\(01\)02177-2](https://doi.org/10.1016/s0169-5347(01)02177-2).

Turner, E., D. J. Jacobson, and J. W. Taylor. 2010. “Reinforced Postmating Reproductive Isolation Barriers in *Neurospora*, an Ascomycete Microfungus.” *Journal of Evolutionary Biology* 23 (8): 1642–56. <https://doi.org/10.1111/j.1420-9101.2010.02030.x>.

VIA, SARA, and JOAN WEST. 2008. “The Genetic Mosaic Suggests a New Role for Hitchhiking in Ecological Speciation.” *Molecular Ecology* 17 (19): 4334–45. <https://doi.org/10.1111/j.1365-294X.2008.03921.x>.

Wu, Chung-I. 2001. “The Genic View of the Process of Speciation.” *Journal of Evolutionary Biology* 14 (6): 851–65. <https://doi.org/10.1046/j.1420-9101.2001.00335.x>.

Yukilevich, Roman. 2012. “ASYMMETRICAL PATTERNS OF SPECIATION UNIQUELY SUPPORT REINFORCEMENT IN *DROSOPHILA*.” *Evolution* 66 (5): 1430–46.
<https://doi.org/10.1111/j.1558-5646.2011.01534.x>.

Yukilevich, Roman, and John R. True. 2006. “Divergent Outcomes of Reinforcement Speciation: The Relative Importance of Assortative Mating and Migration Modification.” *The American Naturalist* 167 (5): 638–54. <https://doi.org/10.1086/503120>.